

"MÉTODO E SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCÁRIOS".

Refere-se o presente relatório

5. a um método e a um sistema automático para a extração das informações preenchidas em um cheque bancário, visando assim a compensação destes documentos.

Como é do conhecimento geral, o processamento de cheques bancários é feito manualmente, sendo que, a partir de 1940 sobre o sistema bancário foi concentrado um grande número e volume de papéis diversos, que resultou do proporcional crescimento da sociedade.

- Assim sendo, encontrar meios para manipular o crescente número de documentos em papel, tornou-se vital para as instituições bancárias em geral.
- 15.

Em 1950, o primeiro processo automático para o processamento de cheques bancários foi iniciado nos Estados Unidos. Os bancos e as indústrias eletrônicas procuraram um processo padrão que pudesse ser usado em todos os bancos daquele país. O sistema, agora conhecido como MICR, foi desenvolvido pelo Stanford Research Institute, utilizando equipamentos projetados pelo General Electric Laboratory. A fonte que foi desenvolvida é conhecida como E-13B, e era somente uma das muitas fontes e formas de caracteres desenvolvidos especialmente para serem utilizados em máquinas de leitura de alta velocidade. Este sistema continua a ser utilizado atualmente, porém, somente processa as informações impressas de cada cheque, isto é,

- 20.
- 25.

através da leitura da linha MICR, é possível identificar o banco, a agência bancária, o número da conta, o número do cheque, etc. Porém, as informações introduzidas no cheque no momento do seu preenchimento pelo cliente, tais como:

5. valor por extenso, favorecido, local, data, e assinatura, continuam sendo processadas manualmente.

- Nestes termos, o processamento automático de cheques bancários através de computadores é um tópico de pesquisa que vem sendo investigado há vários
10. anos. Atualmente a maioria das pesquisas estão voltadas para o reconhecimento de caracteres manuscritos e a verificação de assinaturas. O reconhecimento automático de caracteres manuscritos e a verificação automática de assinaturas já são utilizados com certa confiabilidade. Diversos sistemas e métodos estão disponíveis. Porém, todos estes sistemas trabalham com uma restrição: os elementos a serem reconhecidos e/ou verificados devem estar isolados, isto é, sem a presença de fundos ilustrativos ou coloridos, informações impressas e ou fatores que possam interferir nas
 15. partes a serem reconhecidas/verificadas.

- Não existe atualmente no mercado, nenhum sistema que faça o processamento de cheques bancários automaticamente. Vários países do mundo tem demonstrado interesse nesta área e desenvolvem estudos na
20. tentativa de resolver o problema do processamento de cheques bancários. Alguns poucos trabalhos apresentam métodos que resolvem partes do problema, mesmo assim, vários problemas básicos ainda estão longe de serem resolvidos.

Um estudo intitulado "An Automatic Reading System For Handwritten Numeral Amounts on French Checks" apresentado na Terceira Conferência Internacional de Análise e Reconhecimento de Documentos, descreve

5. um método para a leitura automática do valor numérico de cheques bancários franceses baseado em um modelo probabilístico de segmentação por reconhecimento. O algoritmo determina regiões de corte e fornece representações espaciais múltiplas.
10. Um estudo intitulado "Automatic Processing of Information on Cheques" apresentado na Conferência Internacional de Sistemas, Homem e Cibernética, descreve os procedimentos envolvidos na localização, extração e processamento das informações contidas em cheques bancários canadenses.

20. Um estudo intitulado "Automatic Extraction of Items From Cheque Images for Payment Recognition", apresentado na Conferência Internacional de Reconhecimento de Padrões, em Viena, Austria 1996, descreve um método para extrair os diferentes itens de informação preenchidos de cheques bancários canadenses.

25. Um trabalho intitulado "Detection of Courtesy Amount Block on Bank Checks", também apresentado na Conferência Internacional de Análise e Reconhecimento de Documentos, em Montreal, Canadá, 1995, apresenta uma técnica para localizar o campo onde o valor numérico do cheque é preenchido. Esta técnica é desenvolvida para ser aplicada em cheques bancários americanos.

Um outro trabalho intitulado "A Microprocessor-Based Optical Character Recognition Check Reader", também apresentado na Conferência Internacional de Análise e Reconhecimento de Documentos, em Montreal, Canadá, 1995, apresenta o projeto de um módulo de leitura dos caracteres MICR, impressos em tinta magnética, presentes na parte inferior de cheques bancários chineses.

Um trabalho intitulado "A Structural Method With Local Refining For Handwritten Character Recognition", também apresentado na Conferência Internacional de Análise e Reconhecimento de Documentos, em Montreal, Canadá, 1995, apresenta um algoritmo dedicado ao reconhecimento do valor numérico em cheques bancários italianos.

Um trabalho intitulado "Extraction of Signature and Seal Imprint from Bank Checks by Using Color Information", também apresentado na Conferência Internacional de Análise e Reconhecimento de Documentos, em Montreal, Canadá, 1995, apresenta um método para extração de assinaturas e carimbos utilizando informações de cor, visando uma verificação automática de cheques bancários japoneses.

Um trabalho intitulado "Symbolic/Neural Recognition of Cursive Amounts on Bank Cheques", também apresentado na Conferência Internacional de Análise e Reconhecimento de Documentos, em Montreal, Canadá, 1995, apresenta uma solução simbólica/neural para o reconhecimento de valores manuscritos por extenso presentes em cheques

bancários franceses.

- Um trabalho intitulado "Recognition of Handwritten Words Using Stochastic Models" também apresentado na Conferência Internacional de Análise e Reconhecimento de Documentos, em Montreal, Canadá, 1995, apresenta um método para o reconhecimento global de um pequeno conjunto de palavras baseado em um estágio de pseudo-segmentação que introduz pontos de referência.
- 5.

- Uma pesquisa realizada no final de 1995 pela Faculdade de Economia e Administração da Universidade de São Paulo, (FEA/USP), com apoio da Federação Brasileira das Associações de Bancos (FEBRABAN) demonstrou que 52% dos entrevistados preferem os cheques aos cartões magnéticos, sendo que, 65% dos entrevistados disseram não se importar em preencher cheques diariamente para o pagamento de suas contas.
- 10.
- 15.

- Pesquisa realizada pelo Instituto Gallup nos Estados Unidos, patrocinada pela Associação dos Negociantes Financeiros (FSA) demonstrou que a maioria dos americanos, cerca de 83%, ainda consideram o cheque como a forma de pagamento mais conveniente de suas contas diárias. O hábito prevalece apesar dos pesados investimentos em automação e campanhas publicitárias pelo uso do cartão magnético.
- 20.

- As instituições financeiras tem grande interesse em sistemas que sejam capazes de processar cheques bancários com segurança. Atualmente o processo de compensação de cheques bancários é feito de maneira semi-
- 25.

automática com a presença de operadores que lêem o valor preenchido em cada cheque e digitam os valores lidos em terminais. A verificação de assinaturas é feita de maneira inteiramente manual através de comparação visual entre a

5. assinatura presente no cheque e a contida em um cartão de assinaturas em poder do banco. Porém, a verificação de assinatura não é feita em todos os cheques, somente nos que estejam acima de um determinado valor são efetivamente submetidos à etapa de verificação.

10. A presente invenção tem como objetivo fornecer uma alternativa para a compensação de cheques bancários sem a intervenção humana, em total contraste com as técnicas atuais onde os cheques bancários são processados de forma predominantemente manual, ocasionando

15. um alto custo operacional para a instituição bancária, e permitindo a ocorrência de condições de ineficiência.

Um outro objetivo desta patente de invenção reside em prover um novo método e um igualmente novo sistema para o reconhecimento de cheques bancários, visando extrair as informações introduzidas pelo cliente no momento do preenchimento, bem como identificando as demais informações impressas, permitindo assim fornecer estas informações para outros sistemas dedicados ao reconhecimento ou verificação de informações manuscritas.

20.

25. Os objetivos acima mencionados, fundamentam-se no fato de que atualmente, um grande volume de dinheiro circula todos os dias na forma de cheques bancários manuscritos ou preenchidos à máquina. Quando um che-

que bancário é depositado para crédito em conta, o banco que o recebe está interessado, principalmente, nos seguintes campos: número de compensação, número do banco, número da agência, número da conta, valor do cheque e assinatura, sendo que os outros campos são normalmente ignorados na rotina de compensação dos cheques.

Com a utilização da presente invenção, os cheques serão colocados em um digitalizador óptico associado a uma leitora MICR. As informações lidas por estes dois dispositivos são passadas a um computador digital que faz o processamento das informações, fornecendo como resultado diversas informações isoladas que podem ser então passadas para outros sistemas/módulos responsáveis pelo conhecimento/verificação.

Com isto, todo o processo de compensação de um cheque bancário pode ser feito automaticamente, sem a presença de digitadores ou operadores, resultando em uma substancial redução de custos operacionais, aumento de eficiência e segurança, bem como redução do tempo para a compensação de documentos.

O presente sistema digitaliza um documento, tal como um cheque bancário gerando uma imagem digital e em seguida processa esta imagem identificando os componentes e extraíndo as informações que são de interesse, permitindo que outros sistemas possam processar automaticamente os diversos campos de um cheque, como: sistemas para o reconhecimento de dígitos, sistemas para o reconhecimento de palavras e sistemas de verificação de assi-

naturas.

A seguir, a patente em questão será pormenorizadamente descrita com referência aos desenhos abaixo relacionados, nos quais:

5. a figura 1 ilustra um diagrama de blocos referente às operações básicas de um sistema de processamento de cheques bancários;
- a figura 2 ilustra de forma esquemática os elementos básicos do sistema ora proposto;
10. a figura 3 ilustra um diagrama de blocos funcional do sistema proposto;
- a figura 4 ilustra um fluxograma para o ajuste do ângulo de inclinação do documento;
- a figura 5 ilustra um fluxograma para o ajuste de
15. posição vertical do documento;
- a figura 6 ilustra um fluxograma para o ajuste de posição horizontal do documento;
- a figura 7 ilustra uma representação exemplificativa das informações disponíveis na base de
20. dados;
- a figura 8 ilustra uma vista de uma representação exemplificativa do padrão de fundo de um cheque bancário;
- a figura 9 ilustra um exemplo de uma imagem contendo as informações impressas geradas;
25. a figura 10 ilustra um fluxograma para a geração das informações impressas;
- a figura 11 ilustra um fluxograma para a subtração

- entre imagens;
- a figura 12 ilustra um fluxograma para a limiarização da imagem;
- a figura 13 ilustra um fluxograma para a filtragem da imagem;
5. a figura 14 ilustra um fluxograma para a extração de informações;
- a figura 15 ilustra um fluxograma para a operação do sistema;
10. a figura 16 ilustra um exemplo da imagem final resultante do processamento; e
- a figura 17 ilustra um exemplo dos diferentes itens extraídos;

O dimensionamento e a implementação de um sistema automático para o processamento de cheques bancários exige soluções para os seguintes problemas: aquisição dos dados, identificação/reconhecimento do documento, processamento de imagem/extração de informações e processamento/validação das informações.

20. A figura 1 que integra a patente em questão, ilustra de forma esquemática um sistema típico de processamento de cheques bancários. O método proposto apresenta soluções para os três primeiros blocos do referido diagrama, ou seja, Digitalização, Identificação e
25. Extração.

Na figura 2 é mostrada uma configuração típica do presente invento que consiste de um digitalizador óptico 1, uma leitora MICR 2 (Magnetic Ink

Character Recognition), padrão CMC-7 (Caracteres Magnéticos Codificados em 7 barras) e um microcomputador 3.

- O digitalizador óptico 1 consiste de um dispositivo de leitura óptica que faz a varredura do cheque bancário e gera uma matriz de números no intervalo 0 a 255 representando as intensidades da luz como um conjunto de pontos discretos. O digitalizador óptico utilizado é o "Scanjet IIIp", fabricado pela companhia Hewlett Packard, e é ajustado para fornecer uma imagem digitalizada com 200 pontos por polegada e 256 níveis de cinza. A leitora MICR, identificada na figura 2, pela referência numérica 2, consiste em um dispositivo de leitura óptica que lê os caracteres impressos com tinta magnética (MICR), padrão CMC-7, impressos na parte inferior do cheque. A leitora MICR utilizada é a LCB 2100 fabricada pela empresa Itaotec. O scanner é ligado ao microcomputador através de uma interface SCSI conectada ao barramento ISA, enquanto que a leitora MICR é ligada através da porta serial padrão RS-232.
- Os blocos restantes, responsáveis pelo processamento das informações digitalizadas, são implementados através de um programa de computador. Este programa é constituído por vários blocos, conforme apresentado no diagrama da figura 3.
- O primeiro bloco faz o ajuste da posição da imagem digitalizada. A inclinação é avaliada e se necessário ela é rotacionada até a inclinação zero a ser obtida.

A figura 4 apresenta o fluxograma do algoritmo de ajuste do ângulo de inclinação. Alguns parâmetros são utilizados: *Nmáx* é o número de divisões (janelas) verticais ao longo do eixo horizontal, em que o

5. cheque será dividido; *N* refere-se ao número da janela que será processada; *MáxPPh* é o ponto de máximo do perfil de projeção horizontal e *WnRef* é a coordenada *y* onde este máximo ocorre; *WRef* é o ponto de referência a ser calculado, para o qual todas as janelas serão deslocadas; *ImShift*

10. refere-se ao valor do deslocamento a ser aplicado à janela para corrigir sua posição. Em seguida as posições horizontal e vertical são avaliadas e corrigidas se necessário.

Nas figuras 5 e 6 é apresentado o fluxograma para os algoritmos de ajuste da posição

15. vertical e horizontal respectivamente.

Para o ajuste da posição vertical são utilizados os seguintes parâmetros: *MáxDigPPh* é o ponto de máximo do perfil de projeção horizontal da imagem digitalizada, enquanto *PosMaxDigPPh* é a coordenada *y* onde

20. este máximo ocorre; *MáxGerPPh* é o ponto de máximo do perfil de projeção horizontal da imagem gerada, enquanto *PosMaxGerPPh* é a coordenada *y* onde este máximo ocorre; *VertShift* refere-se ao valor do deslocamento a ser aplicado à janela para corrigir sua posição.

25. Para o ajuste de posição horizontal são utilizados os seguintes parâmetros: *H1* que representa a primeira linha horizontal, *H2* que representa a segunda linha horizontal, *L11* que representa a décima pri-

meira linha vertical e finalmente, *L15* que representa a décima quinta linha horizontal.

É selecionada a região limitada por estas quatro linhas, tanto na imagem digitalizada

5. (*Dig*), quanto na imagem gerada (*Ger*).

É traçado o perfil de projeção vertical para a imagem digitalizada (*DigPPv*) e o valor de máximo (*MáxDigPPv*) é calculado, bem como a posição onde ele ocorre (*PosMáxDigPPv*), é também determinada. O mesmo é fei-

10. to para a imagem gerada (*Ger*), sendo encontrados então os parâmetros *GerPPv*, *MáxGerPPv* e *PosMáxGerPPv*. Em seguida é determinada a diferença entre a posição das duas imagens, dado por *HorShift*. A imagem digitalizada (*Dig*) é então deslocada *Horshift* pixels, ficando, assim, ajustada em relação
15. à imagem gerada.

O próximo bloco é responsável pela geração de uma imagem que contenha todas as informações presentes em um cheque. Este bloco faz uso das informações armazenadas em uma base de dados, onde estão dispo-

20. níveis as informações necessárias para a geração da imagem com as informações impressas, como: nome do cliente, número da conta, número do cheque, número do banco, número da agência, número de compensação, dígitos de compensação, nome do banco, nome da agência, posição e dimensão das linhas
25. horizontais e verticais, etc.

Além disso, devemos ter a posição de cada um destes elementos dentro do cheque e o tipo de fonte, tamanho, estilo, etc, usado em cada elemento.

Na figura 7 é apresentado um exemplo dos elementos presentes na base de dados. Na base de dados é armazenada também uma amostra do padrão de fundo de cada cheque que será processado.

5. Na figura 8 é apresentado um exemplo do padrão de fundo de um cheque bancário. O módulo de geração de informações utiliza estas informações da base de dados para gerar a imagem contendo somente as informações impressas presente em um cheque.

10. Esta imagem é similar a apresentada na figura 9. Na figura 10 é apresentado o fluxograma para o algoritmo de geração da imagem com as informações impressas.

Após a identificação de alguns

15. parâmetros dos cheques como *Comp*, *BANCO*, *AG*, *NC* e *NCh*, através da leitura da linha MICR, estas informações são confrontadas com as existentes na base de dados fornecendo como resposta se o cheque pode ser processado ou não.

Em caso positivo, é então alocado na memória do microcomputador (referência 3 da figura 2) um espaço para a criação de uma imagem do tipo *Bitmap*, então, a base de dados é novamente acessada e é feita a leitura de todas as informações necessárias para a geração da imagem contendo todas as informações impressas presentes

20. no cheque.

O bloco seguinte é responsável pela subtração entre a imagem do cheque com a posição ajustada ($IC(x,y)$), a imagem gerada com as informações impres

sas ($IG(x,y)$) e a amostra do padrão de fundo ($IF(x,y)$). Este bloco trata as imagens como matrizes que são subtraídas ponto a ponto até a última coordenada (x,y).

Na figura 11 é apresentado o

5. fluxograma para as operações realizadas pelo algoritmo de subtração. O próximo bloco (bloco de Thresholding da figura 3) faz a limiarização da imagem resultante da operação de subtração. Este bloco faz a redução dos 256 níveis de cinza da imagem para apenas dois níveis.

10. Para a imagem $IR(x,y)$ é escolhido um valor para o limiar Th de acordo com o banco emissor do cheque.

A partir das coordenadas (x,y)

= (1,1) a imagem é testada ponto a ponto até o seu final.

15. Caso o valor do pixel na coordenada seja maior do que Th o pixel irá assumir o valor 0 (preto), caso contrário, assumirá o valor 1 (branco).

Na figura 12 é apresentado o

fluxograma que executa as operações deste bloco. O último

20. bloco de processamento da imagem faz uma filtragem da imagem resultante. Esta imagem pode conter algum ruído. Este bloco executa uma operação morfológica de dilatação.

Na figura 13 é apresentado o

fluxograma das operações realizadas por este bloco. Inicialmente o algoritmo lê a imagem $IR(x,y)$ e a máscara de

25. filtragem $MS(x,y)$. Esta máscara de filtragem é de dimensão 2 x 2 e é dada por:

$$MS(x,y) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

A partir das coordenadas (x,y)

5. = (1,1) a imagem é testada ponto a ponto até o seu final. Se o pixel da máscara é diferente de zero então é feita a subtração entre os pixels da imagem e da máscara para esta posição.

Se o resultado desta operação

10. for diferente de zero, então, é feito uma operação de máximo entre os elementos da máscara e os da vizinhança do pixel sob operação da imagem. Finalmente, o resultado deste processamento será a imagem do cheque contendo somente as informações preenchidas. Os demais componentes do cheque,
15. como o padrão de fundo e as informações impressas, foram eliminados. Por último os diversos itens de informação são separados entre si pelo bloco de extração de informações. O fluxograma do algoritmo que implementa este bloco é apresentado na figura 14.

20. O sistema proposto inclui pelo menos um terminal que consiste de um microcomputador 3, um digitalizador óptico 1 e uma leitora MICR 2.

- No microcomputador 3, deve estar disponível um banco de dados com todas as informações
25. necessárias sobre o banco, agência e clientes, além é claro, dos programas que implementam os algoritmos de processamento apresentados.

A figura 15 mostra um diagrama

de blocos com cada etapa da operação do sistema. Um cheque é introduzido ao sistema inicialmente através da leitora MICR 2. A leitura dos caracteres padrão CMC-7 da linha MICR, impressos com tinta magnética, fornecem ao sistema in-

5. informações que possibilitam a identificação do banco, da agência, do número de série do cheque, do número da conta, do código de compensação, dos dígitos verificadores e dos códigos de tipificação.

Estas informações são passadas

10. à base de dados que interpreta estes códigos. A base de dados identifica o cheque e todas as informações a ele relacionadas, sendo, então, passadas dois tipos de informações ao bloco seguinte; uma imagem com o padrão de fundo do cheque e uma seqüência de caracteres e instruções. A imagem
15. do padrão de fundo do cheque é passada diretamente ao bloco de subtração, enquanto as seqüências de caracteres e as instruções são passadas ao bloco de geração de informações impressas.

O bloco de geração de informa-

20. ções impressas é responsável pela geração de uma imagem que contém todas as informações presentes no cheque bancário que está sendo processado, incluindo, as seqüências de caracteres, as linhas verticais e horizontais, os símbolos e objetos geométricos.

25. Na figura 9 é ilustrado um exemplo genérico de uma imagem com as informações impressas. A imagem gerada contendo somente as informações impressas é também passada ao bloco de subtração.

Paralelamente a estes processos, logo após o cheque bancário ter passado pela leitora MICR 2, ele é levado a um outro dispositivo de aquisição, desta vez, um digitalizador óptico 1.

5. Este digitalizador ou scanner óptico faz uma varredura no cheque e gera uma matriz de números representando as intensidades de luz como um conjunto de pontos discretos. Assim, tem-se uma imagem digital em um formato compatível com o processamento computacional.

10. A seguir esta imagem é levada ao bloco de ajuste de posição responsável pela normalização da imagem, ou seja, as linhas verticais e horizontais impressas sobre o cheque devem ser perpendiculares às bordas do papel.

15. A imagem resultante do processo de ajuste de posição é, então, passada ao bloco de subtração. Como foi visto, o bloco de subtração recebe três imagens: *o padrão de fundo do cheque, a imagem com as informações impressas geradas e a imagem digitalizada norma*

20. *lizada do cheque.* Este bloco, como o próprio nome sugere, faz a operação de subtração entre as três imagens. Inicialmente a imagem digitalizada normalizada do cheque é subtraída da imagem do padrão de fundo. Em seguida a imagem resultante desta operação é novamente subtraída, mas desta

25. vez da imagem com as informações impressas gerada. Como resultado tem-se uma imagem que contém somente as informações que foram introduzidas no momento do preenchimento do cheque.

A imagem resultante do processo de subtração é representada por 256 níveis de cinza, porém, apenas dois níveis são suficientes, pois teremos um fundo branco e as informações preenchidas em preto.

5. Assim, esta imagem é submetida a uma operação de "*thresholding*" que faz o mapeamento dos 256 níveis para 2 níveis, fornecendo, assim uma imagem binária.

10. A imagem resultante do processo de subtração pode apresentar alguns pontos distribuídos ao longo de toda sua área. Estes pontos, chamados de ruídos, são devidos a imperfeições no ajuste de posição e pequenas diferenças dos níveis de cinza entre as imagens subtraídas.

15. Para minimizar este efeito, a imagem passa por um bloco de filtragem, onde uma operação morfológica de dilatação com um elemento estruturante é realizada. Na figura 16 tem-se um exemplo da imagem resultante destas etapas de processamento.

20. Finalmente, a última operação consiste em separar os diferentes itens de informação presentes na imagem final, que são: o valor numérico, o valor por extenso, o nome do favorecido, o local e data de emissão, e a assinatura. O bloco de extração de informações
25. utilizando um conhecimento a priori da provável posição de aparecimento de cada item, divide a imagem resultante em diferentes regiões, separando, assim os diferentes itens, como mostrado na figura 17.

A invenção ora tratada é o primeiro sistema para a extração automática das informações preenchidas pelo cliente em cheques bancários. O sistema é simples e de baixo custo, visto que necessita apenas de um

5. microcomputador pessoal, um digitalizador óptico e uma leitora MICR.

A utilização deste sistema permite a manipulação de cheques bancários sem a intervenção humana, de maneira inteiramente automática. Se integrado com outros sistemas automáticos, como por exemplo, reconhecimento de numerais, reconhecimento de palavras e verificação de assinaturas, pode realizar a tarefa de compensação de cheques bancários automaticamente. O sistema oferece uma resposta aceitável menor do que 1 minuto, para o processamento de cada cheque, o que é adequado para os serviços bancários. O método permite a manipulação de cheques bancários sem a presença de digitadores, diminuindo erros e aumentando a segurança nas transações financeiras envolvendo cheques bancários. Além disso, a aplicação deste sistema pode reduzir significativamente a circulação de documentos em papel entre agências e entre bancos, que pode ser feito de maneira digital através de uma rede de comunicações.

REIVINDICAÇÕES

1. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", caracterizado pelo fato de compreender a introdução no sistema de um cheque bancário, através de uma leitora MIRC (2) para a leitura dos caracteres padrão CMC-07 da linha MICR, impressos com tinta magnética.

2. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 1, caracterizado pelo fato de que a leitura dos caracteres padrão CMC-07 da linha MICR, impressos com tinta magnética, fornecem ao sistema informações que possibilitam a identificação do banco, da agência, do número de série do cheque, do número da conta, do código de compensação, dos dígitos verificadores e do código de tipificação.

3. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 1, caracterizado pelo fato de que as informações obtidas pela leitura dos caracteres padrão CMC-07 são passadas à uma base de dados, que interpreta estes códigos.

4. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 3, caracterizado pelo fato de que a base de dados identifica o cheque e todas as informações a ele relacionadas.

5. "MÉTODO PARA EXTRAÇÃO AUTO-

MATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 1 e 4, caracterizado pelo fato de que com base na identificação procedida pela base de dados, são obtidos dois tipos de informação, sendo uma imagem com

5. o padrão de fundo do cheque e uma seqüência de caracteres e instruções.

6. "MÉTODO PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 1 e 6, caracterizado pelo fato de

10. que a imagem do padrão de fundo do cheque é passada diretamente a um bloco de subtração, enquanto as seqüências de caracteres e as instruções são passadas ao bloco de geração de informações impressas.

7. "MÉTODO PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 6, caracterizado pelo fato de que o bloco de geração de informações impressas é responsável pela geração de uma imagem que contém todas as informações impressas presentes no cheque bancário que esta sendo pro-

15. cessado, incluindo as seqüências de caracteres, as linhas verticais e horizontais, os símbolos e objetos geométricos.

20.

8. "MÉTODO PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 7, caracterizado pelo fato de que

25. a imagem gerada pelo bloco de geração contendo somente as informações impressas é passada também ao bloco de subtração.

9. "MÉTODO PARA EXTRAÇÃO AUTO-

MATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 1 e 8, caracterizado pelo fato de que paralelamente às etapas precedentes, após o cheque bancário ter passado pela leitora MICR (2), este é levado a um

5. outro dispositivo de aquisição, que compreende um digitalizador óptico (1), que faz a varredura no cheque bancário e gera uma matriz de números representando as intensidades da luz como um conjunto de pontos discretos.

10. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 9, caracterizado pelo fato de que após o processamento do cheque bancário no digitalizador óptico (1), obtém-se uma imagem digital em formato compatível com o processamento computacional.

15. 11. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 10, caracterizado pelo fato de que a informação obtida pelo digitalizador óptico (1), é levada a um bloco de ajuste de posição, responsável pela

20. normalização da imagem, de modo a que a referida imagem seja ajustada no tocante às suas linhas verticais e horizontais impressas, que devem ser perpendiculares às bordas do papel.

25. 12. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 11, caracterizado pelo fato de que a imagem resultante da etapa de ajuste de posição é, então, passada ao bloco de subtração.

13. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 6, 8 e 12, caracterizado pelo fato de que o bloco de subtração recebe três imagens, ou seja, o padrão de fundo do cheque, a imagem com as informações impressas geradas e a imagem digitalizada normalizada do cheque.

14. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 6, 8 e 12, caracterizado pelo fato de que o bloco de subtração faz uma operação de subtração entre as três imagens recebidas.

15. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 14, caracterizado pelo fato de que a subtração é realizada entre a imagem original do cheque, uma amostra do padrão de fundo do cheque e uma imagem que contém as informações impressas geradas pelo bloco de geração de informações impressas.

16. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 15, caracterizado pelo fato de que a imagem resultante do processo de subtração é representada por 256 níveis de cinza.

17. "MÉTODO PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 16, caracterizado pelo fato de

por 256 níveis de cinza, é submetida a uma operação de "thresholding", que faz o mapeamento dos mencionados 256 níveis de cinza, para dois níveis.

18."MÉTODO PARA EXTRAÇÃO AUTO-

5. MATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 17, caracterizado pelo fato de que do mapeamento levado a efeito sobre a imagem obtida do processo de subtração resulta uma imagem binária.

19."MÉTODO PARA EXTRAÇÃO AUTO-

10. MATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 18, caracterizado pelo fato de compreender uma etapa ou bloco de filtragem, que visa eliminar ou minimizar o efeito de ruído que a imagem binária, obtida após a operação de "thresholding", pode apresentar.

15. 20."MÉTODO PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 19, caracterizado pelo fato de que a etapa ou bloco de filtragem compreende uma operação morfológica de dilatação, com um elemento estruturante.

20. 21."MÉTODO PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 20, caracterizado pelo fato de que após a etapa ou bloco de filtragem é realizada a última operação que consiste em separar os diferentes itens de informação presentes na imagem final, que são: o valor numérico, o valor por extenso, o nome do favorecido, o local e data de emissão, e a assinatura.

22."MÉTODO PARA EXTRAÇÃO AUTO-

MATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", caracterizado pelo fato de prever um bloco ou etapa de extração de informação que compreende o conhecimento prévio da provável posição de aparecimento de cada item de informação no cheque bancário, dito bloco divide a imagem resultante do processamento em diferentes regiões, separando, assim, os diferentes itens.

23. "SISTEMA PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", caracterizado pelo fato de compreender o um digitalizador óptico (1), uma leitora MICR (2) padrão CMC-07 e um microcomputador (3).

24. "SISTEMA PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de que o digitalizador óptico (1) faz uma varredura do cheque bancário e gera uma matriz de números no intervalo 0 a 255, que representa as intensidades da luz como um conjunto de pontos discretos.

25. "SISTEMA PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de que o digitalizador óptico (1) é ajustado para fornecer uma imagem digitalizada com 200 pontos por polegada e 256 níveis de cinza.

26. "SISTEMA PARA EXTRAÇÃO AUTOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de

que a leitora MICR (2) configura-se como um equipamento de leitura que lê os caracteres impressos com tinta magnética (MICR), padrão CMC-07, impressos na parte inferior dos cheques bancários.

5. 27. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de que o digitalizador óptico (1) é ligado ao microcomputador (3) através de uma interface SCSI, conectada ao barramento
10. ISA.
28. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de que a leitora MICR (2) é ligada ao microcomputador (3)
15. através da porta serial padrão RS-232.
29. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de compreender um algoritmo de ajuste do ângulo de inclinação,
20. que utiliza parâmetros como: $N_{máx}$, que é o número de divisões ou janelas verticais ao longo do eixo horizontal, em que o cheque bancário será dividido; N , que refere-se ao número da janela que será processada; $MáxPPH$, que é o ponto de máximo do perfil de projeção horizontal e $WnRef$, que é a
25. coordenada y onde este máximo ocorre; $WRef$ que é o ponto de referência a ser calculado, para o qual todas as janelas serão deslocadas; $ImShift$ que refere-se ao valor do deslocamento a ser aplicado à janela para corrigir sua posição.

30. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de compreender algoritmos de ajuste da posição vertical e horizontal.

31. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 30, caracterizado pelo fato de que o algoritmo para o ajuste da posição vertical se utiliza dos seguintes parâmetros: *MáxDigPPh*, que representa o ponto de máximo do perfil de projeção horizontal da imagem digitalizada; *PosMáxDigPPh*, que representa a coordenada de y, onde este máximo ocorre; *MáxGerPPh*, que representa o ponto de máximo do perfil de projeção horizontal da imagem gerada; *PosMáxGerPPh*, que representa a coordenada de y, onde este máximo ocorre; *VertShift*, que refere-se ao valor do deslocamento a ser aplicada à janela para corrigir sua posição.

32. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 30, caracterizado pelo fato de que o algoritmo para ajuste da posição horizontal se utiliza dos seguintes parâmetros: *H1*, que representa a primeira linha horizontal; *H2* que representa a segunda linha horizontal; *L11*, que representa a décima primeira linha vertical e *L15* que representa a décima quinta linha horizontal.

33. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS",

- segundo o reivindicado em 32, caracterizado pelo fato de que com base nos parâmetros $H1$, $H2$, $L11$ e $L15$, é selecionada a região limitada pelas referidas quatro linhas, tanto na imagem digitalizada (Dig), quanto na imagem gerada
5. (Ger), sendo traçado o perfil de projeção vertical para a imagem digitalizada ($DigPPv$) e o valor máximo ($MáxDigPPv$) é calculado, bem como a posição onde ele ocorre ($PosMáxDigPPv$), é também determinada; o mesmo sendo feito para a imagem gerada (Ger), sendo encontrados então os parâmetros
10. $GerPPv$, $MáxGerPPv$ e $PosMáxGerPPv$.

34. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 33, caracterizado pelo fato de ser determinada a diferença entre a posição das imagens,
15. digitalizada e gerada, dada por $Horshift$; a imagem digitalizada (Dig) é então deslocada $Horshift$ pixels, ficando, assim ajustada em relação à imagem gerada.

35. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS",
20. segundo o reivindicado em 23, caracterizado pelo fato de compreender a geração de uma imagem que contenha todas as informações presentes em um cheque bancário, imagem esta que é obtida mediante a utilização das informações armazenadas em uma base de dados, onde estão disponíveis as in-
25. formações necessárias para a geração da imagem com as informações impressas, como: nome do cliente, número da conta, número do cheque, número do banco, número da agência, número de compensação, dígitos de compensação, nome do

banco, nome da agência, posição e dimensão das linhas horizontais e verticais, entre outras.

36. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS",

5. segundo o reivindicado em 35, caracterizado pelo fato de que na base de dados é armazenada também uma amostra do padrão de fundo de cada cheque que será processado.

37. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS",

10. segundo o reivindicado em 23, caracterizado pelo fato de compreender um algoritmo de geração da imagem com as informações impressas, que após a identificação de parâmetros dos cheques bancários tais como: *Comp*, *BANCO*, *AG*, *NC* e *NCh*, através da leitura da linha MICR, estas informações são
15. confrontadas com as existentes na base de dados fornecendo como resposta se o cheque bancário pode ser processado ou não, sendo que, em caso positivo é então alocado na memória do microcomputador (3) um espaço para a criação de uma imagem do tipo *Bitmap*, então, a base de dados é novamente
20. acessada e é feita a leitura de todas as informações necessárias para a geração da imagem contendo todas as informações impressas presentes no cheque bancário.

38. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS",

25. segundo o reivindicado em 23, caracterizado pelo fato de prever um algoritmo de subtração que executa a subtração entre a imagem do cheque bancário com a posição ajustada ($IC(x,y)$), a imagem gerada com as informações impressas

($IG(X,Y)$) e a amostra do padrão de fundo ($IF(x,y)$), dito bloco de subtração trata a imagem como matrizes que são subtraídas ponto a ponto até a última coordenada (x,y).

39. "SISTEMA PARA EXTRAÇÃO AU-

5. TOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de compreender a limiarização da imagem resultante da operação de subtração, onde é feita a redução dos 256 níveis de cinza da imagem para 2 níveis, sendo que, para a imagem $IR(x,y)$
10. é escolhido um valor para o limiar Th de acordo com o banco emissor do cheque bancário; a partir das coordenadas (x,y) = (1,1) a imagem é testada ponto a ponto até o seu final; caso o valor do pixel na coordenada seja maior que Th o pixel irá assumir o valor 0 (preto), sendo que, em caso
15. contrário, assumirá o valor 1 (branco).

40. "SISTEMA PARA EXTRAÇÃO AU-

- TOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS", segundo o reivindicado em 23, caracterizado pelo fato de compreender a filtragem da imagem resultante, mediante a
20. execução de uma operação morfológica de dilatação, dita filtragem inicia-se pela leitura da imagem $IR(x,y)$ e da máscara de filtragem $MS(x,y)$, por um algoritmo, sendo que, a máscara de filtragem é de dimensão de 2 x 2 e é dada por:

$$MS(x,y) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

41. "SISTEMA PARA EXTRAÇÃO AU-

25. TOMATICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS",

segundo o reivindicado em 40, caracterizado pelo fato de que a partir das coordenadas $(x,y) = (1,1)$ a imagem é testada ponto a ponto até o seu final, sendo que, se o pixel da máscara é diferente de zero então é feita a subtração

5. entre os pixels da imagem e da máscara para esta posição; se o resultado desta operação for diferente de zero, então, é feita uma operação de máximo entre os elementos da máscara e os da vizinhança do pixel sob operação da imagem.

42. "SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCARIOS",

10. segundo o reivindicado em 41, caracterizado pelo fato de produzir como resultado a imagem do cheque bancário contendo somente as informações preenchidas, sendo que após a obtenção da imagem do referido cheque contendo as informações

15. preenchidas, estas informações são separadas entre si por blocos de extração de informação, mediante um algoritmo específico.

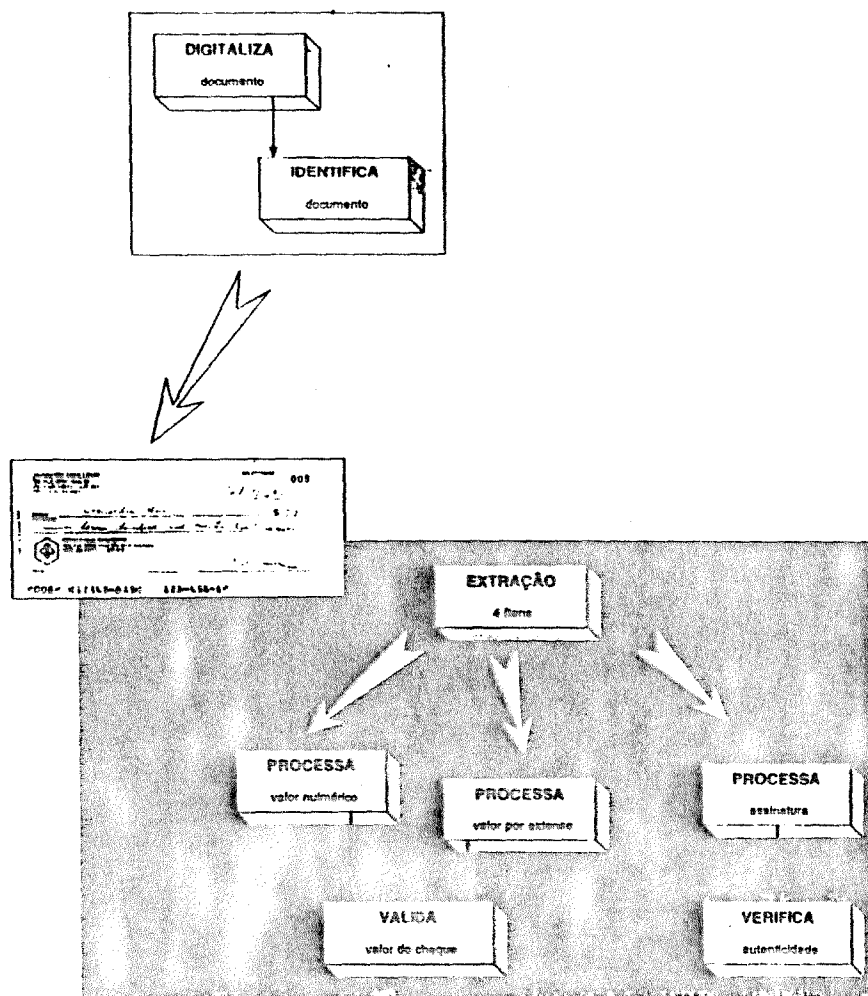
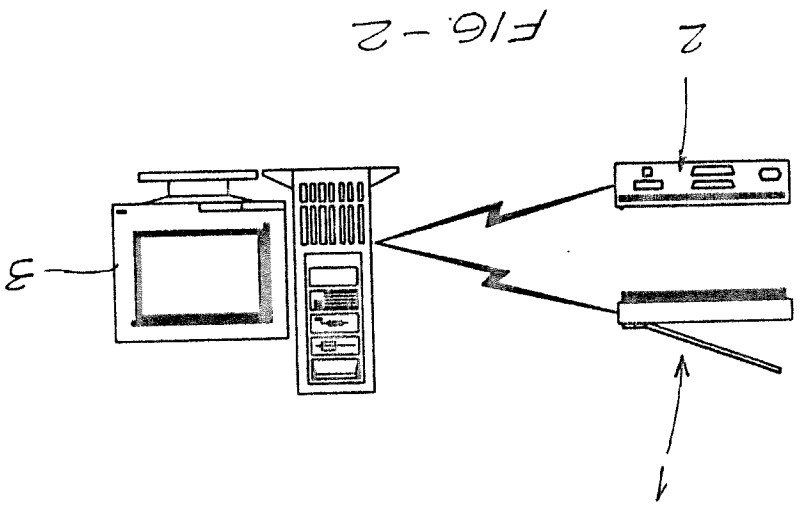


FIG.-1



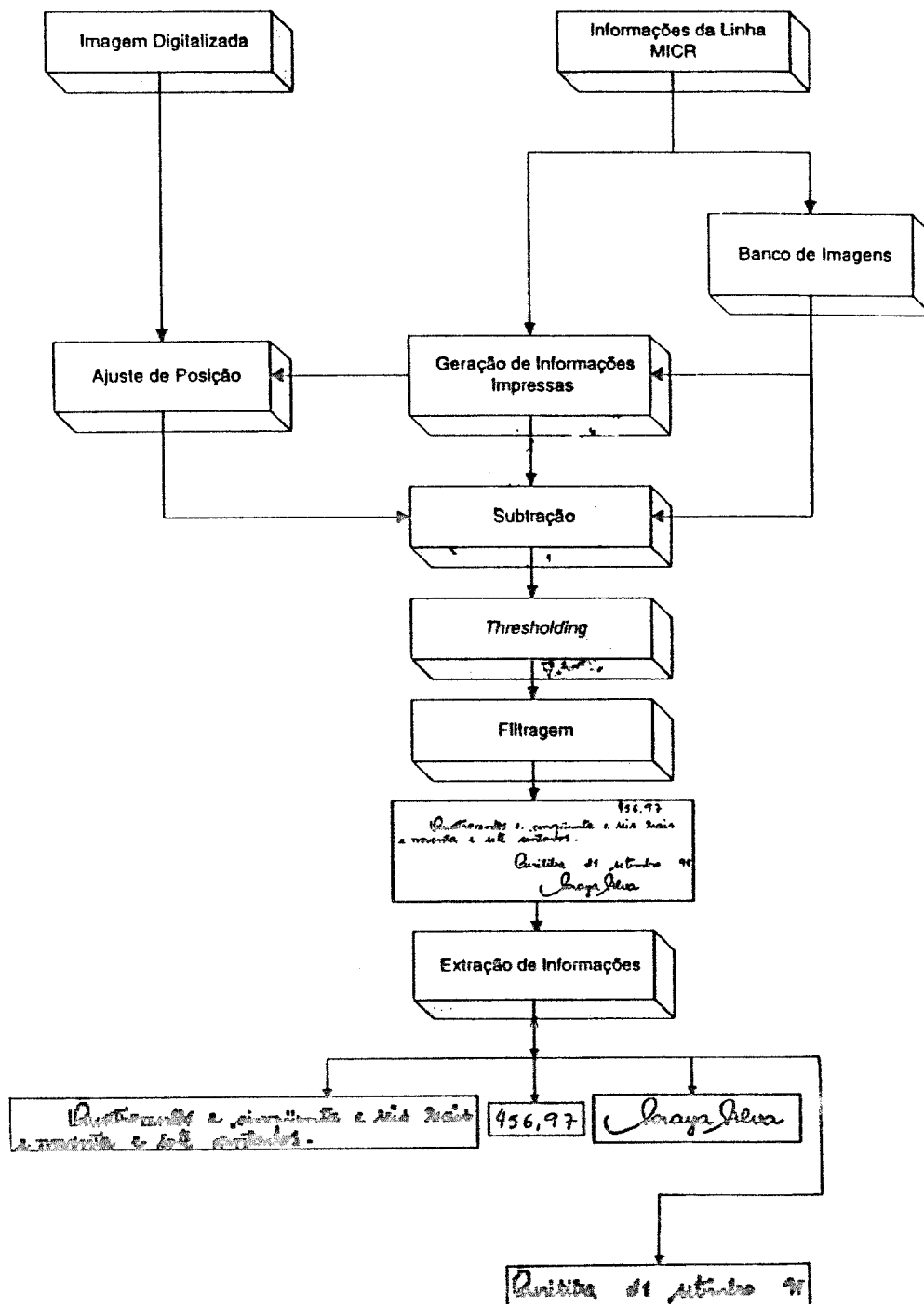


FIG.-3

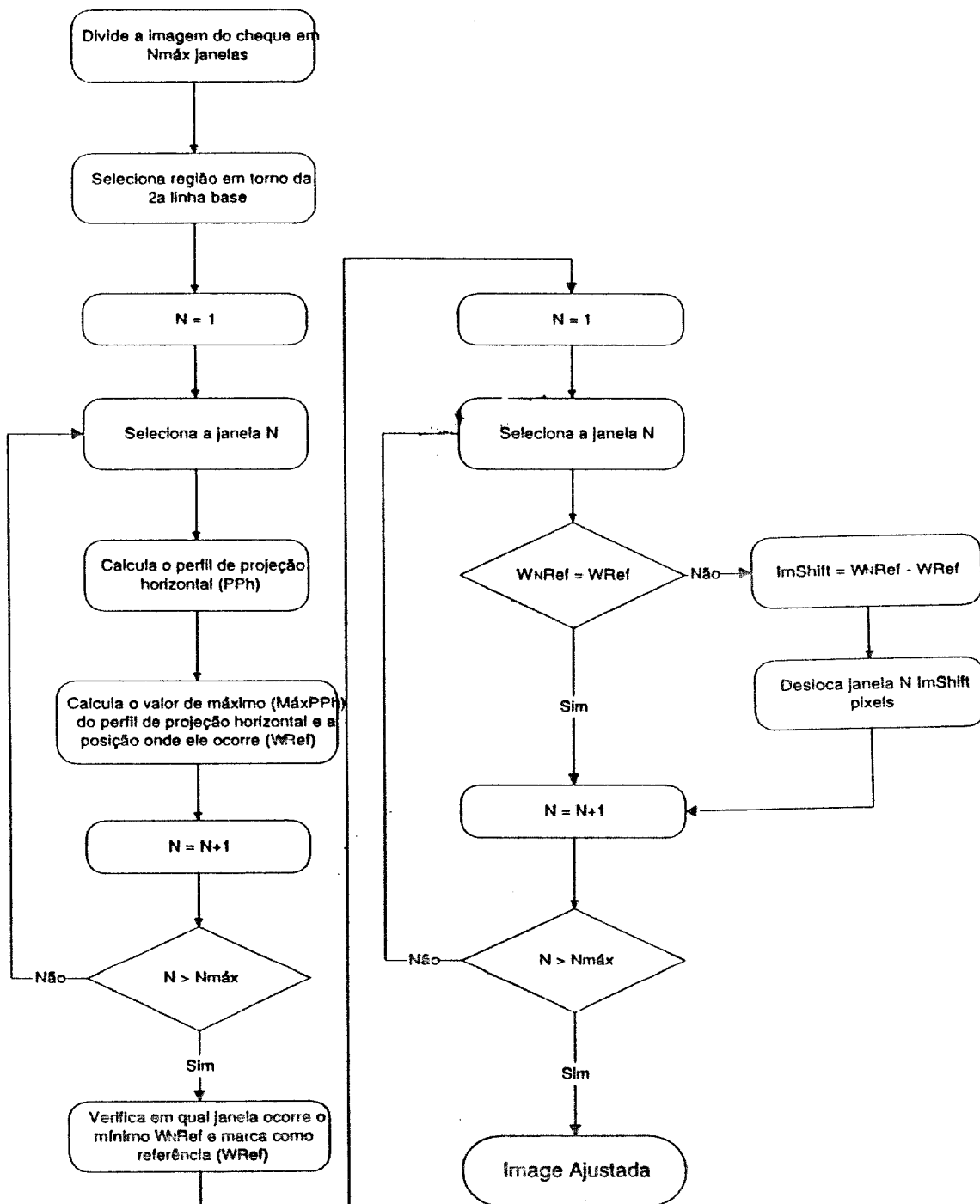


FIG.-4

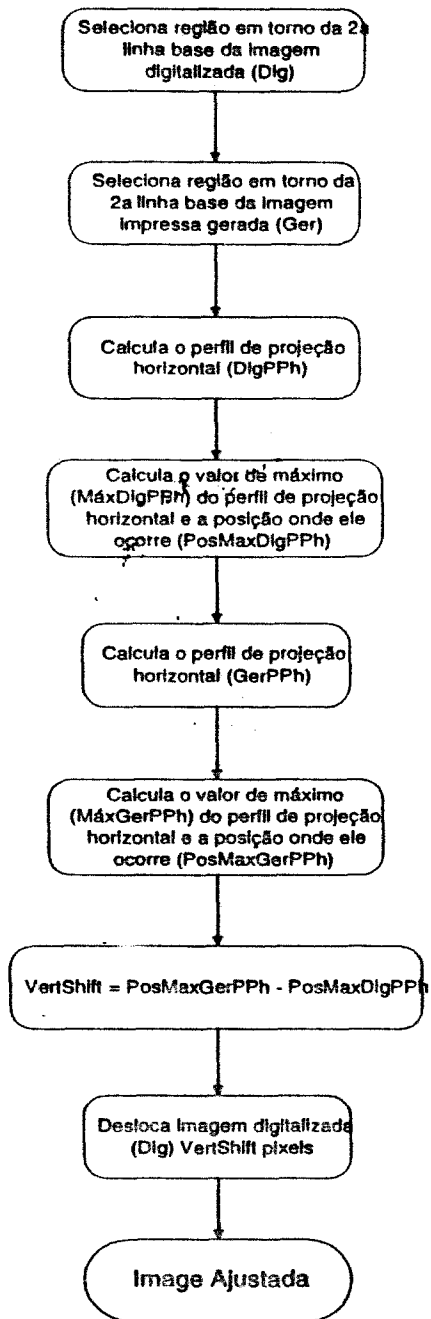


FIG. -5

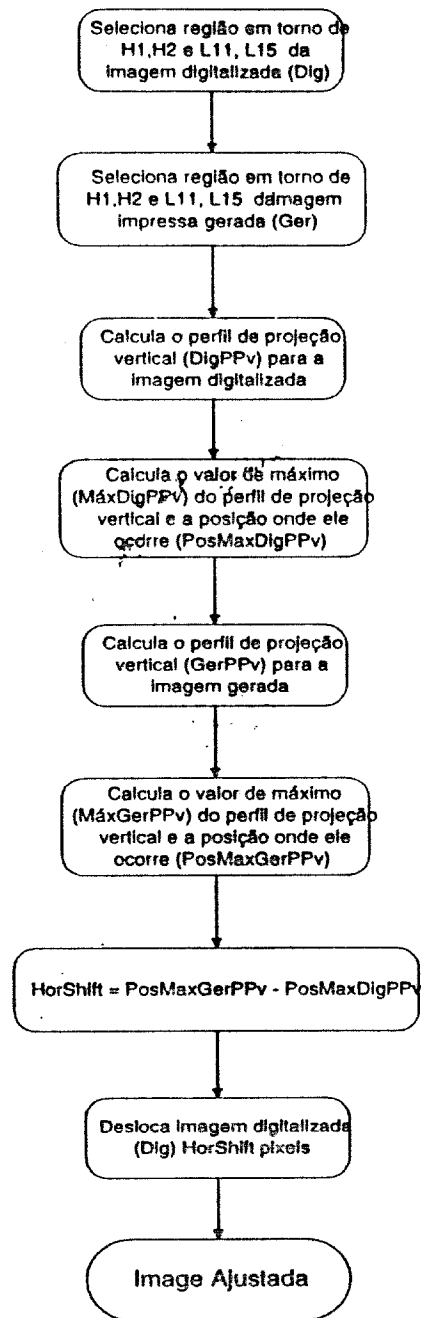


FIG.-6

"001"
"Banco do Brasil S/A", ARB, 0, 0, 15, 120, 460
L1, 12, 1345, 2, 0, 1200
L2, 23, 1345, 2, 0, 1200
L3, 12, 1250, 2, 0, 1130
L4, 620, 1345, 2, 0, 540
L5, 620, 1345, 2, 0, 540
"Comp.", TNR, R, 0, 0, 10, 0, 23, 29
"Banco", TNR, 0, 0, 10, 0, 64, 29
"Agencia", TNS, R, 0, 0, 10, 0, 120, 29
.
.
.
"Numero do cheque", TNR, R, 0, 0, 10, 0, 789, 29
"R\$", TNR, R, 0, 0, 10, 0, 1020, 29
.
.
.
"Pague por", TNR, R, 0, 0, 8, 0, 18, 190
"este cheque a", TNR, R, 0, 0, 8, 0, 18, 200
"quantia de", TNR, R, 0, 0, 8, 0, 18, 210
.
.
.
"Jose Carlos da Silva", TNR, R, 0, 0, 10, 0, 567, 459
"095.903.079-49", TNR, R, 0, 0, 10, 0, 614, 492
"02701855-0162732155-200000025064", MICR, R, 0, 0, 10, 0, 10, 599
"273.215", TNR, R, 0, 0, 10, 789, 29
"000.250-6", TNR, R, 0, 0, 10, 0, 1020, 29
"4", TNR, R, 0, 0, 10, 0, 799, 29
.
.
.

FIG. -7

BANCO RURAL S.A.

F/G-8

Comp	Banco	Agência	C1	Conta	C2	Chaque nº	C3	R\$
009	104	0368	() 6	03000768-0	0 (AAA)	302910	7	

Pague por este cheque a quantia de _____

_____ de _____ de 19____

_____ on a sua ordem

COMENDADOR, PR
R COMENDADOR ARAUJO, 117
CURITIBA - PR

ASSOCIACAO DE ENSINO NOVO ATENEU
76.534.924/0001-30 CAIXA 100

FIG.-9

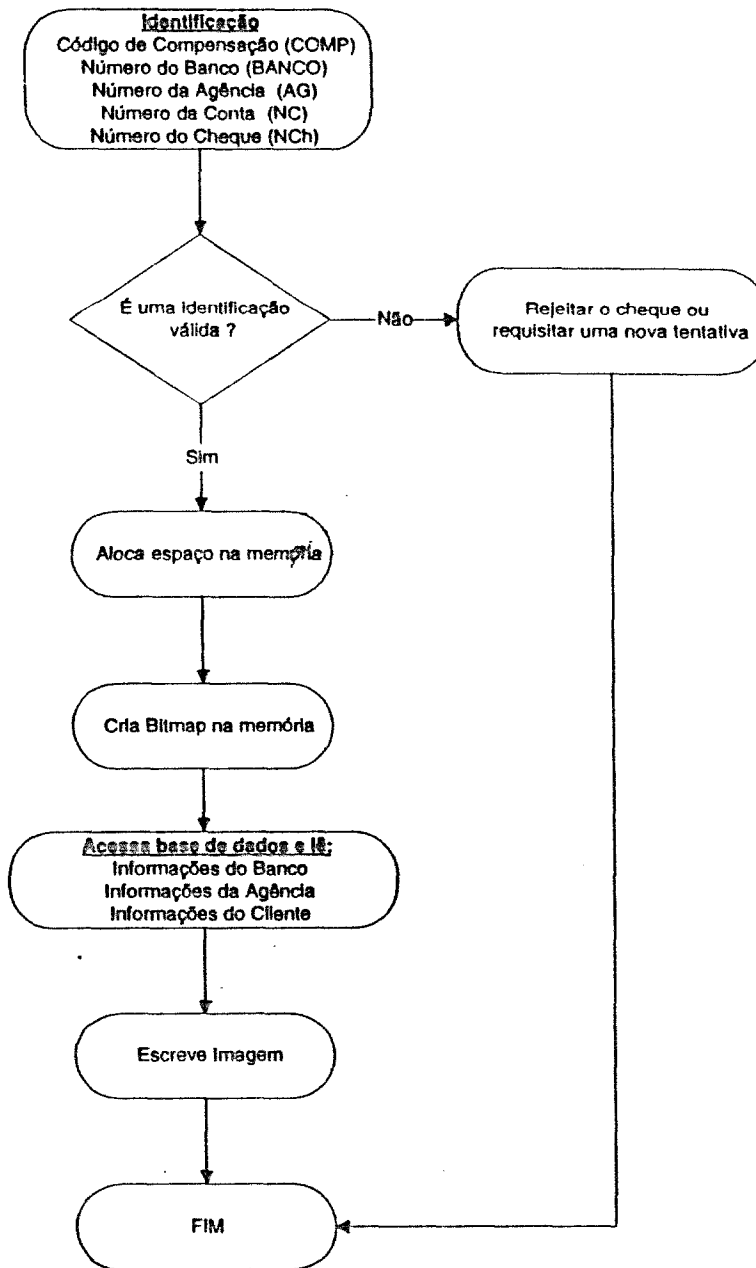


FIG. -10

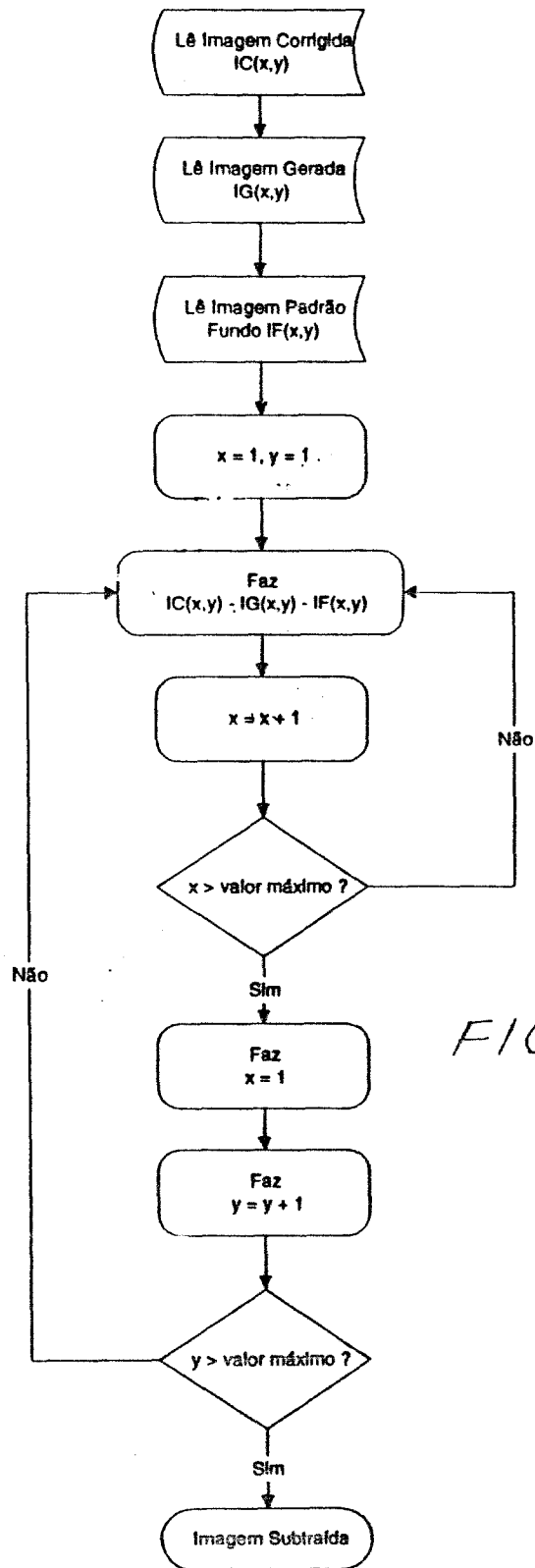


FIG. 11

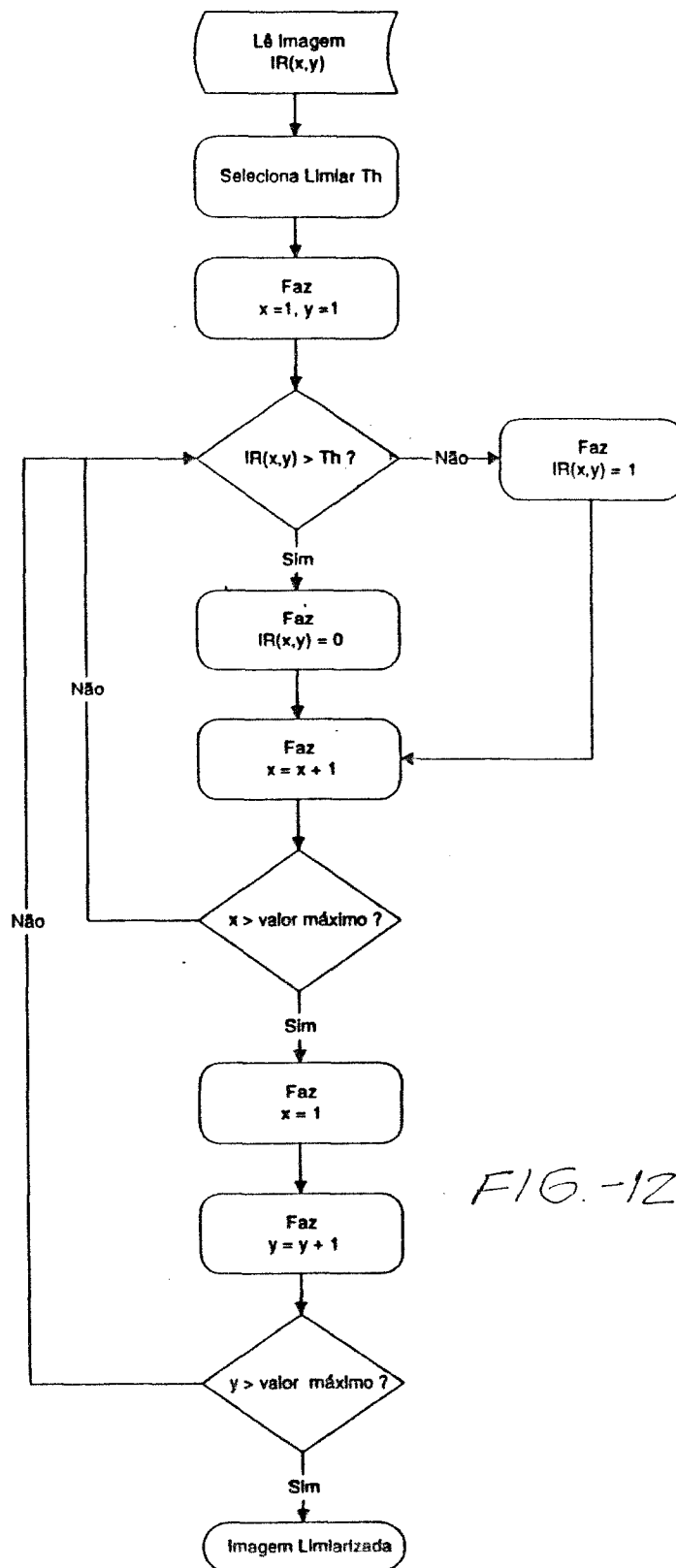
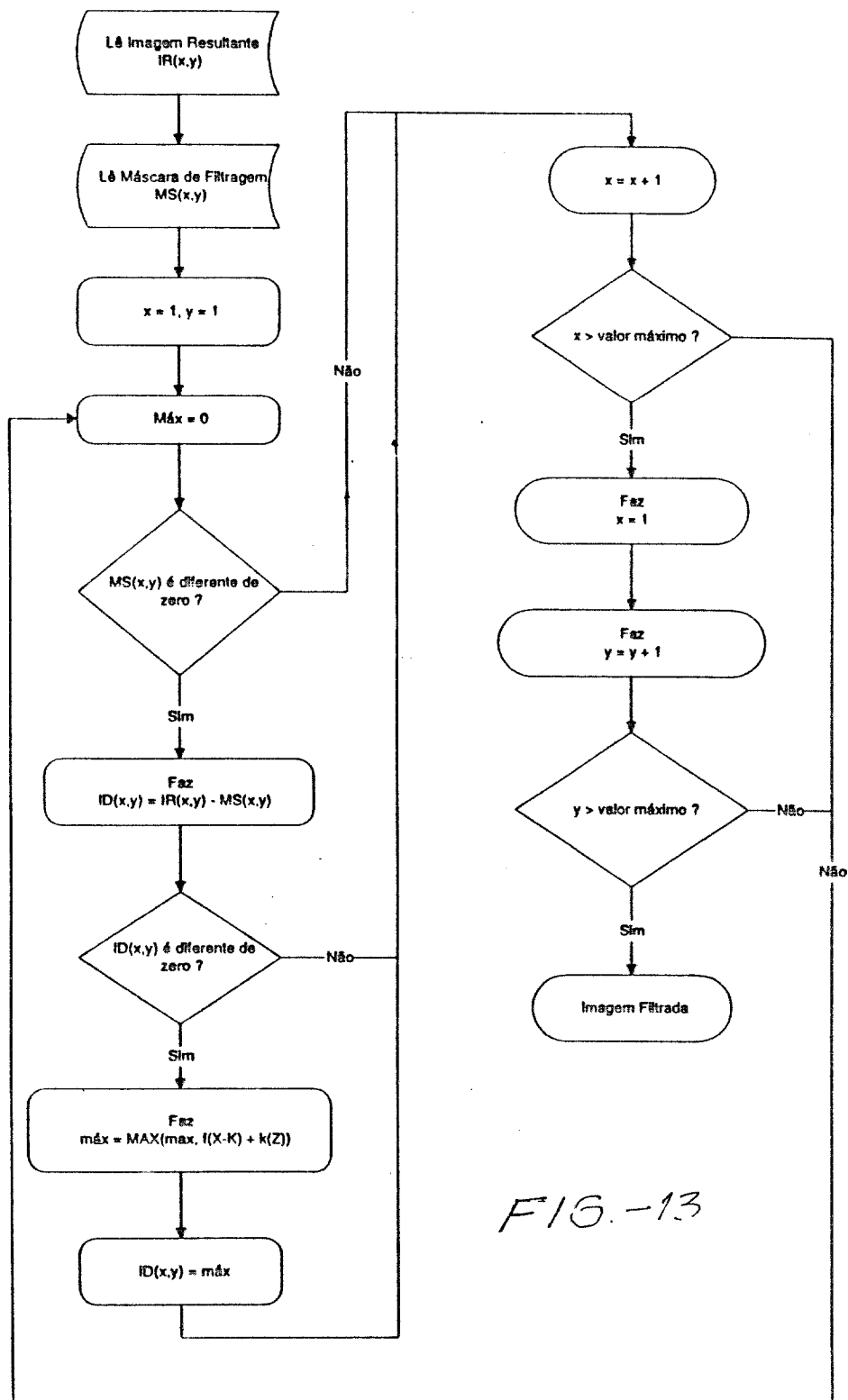


FIG.-12



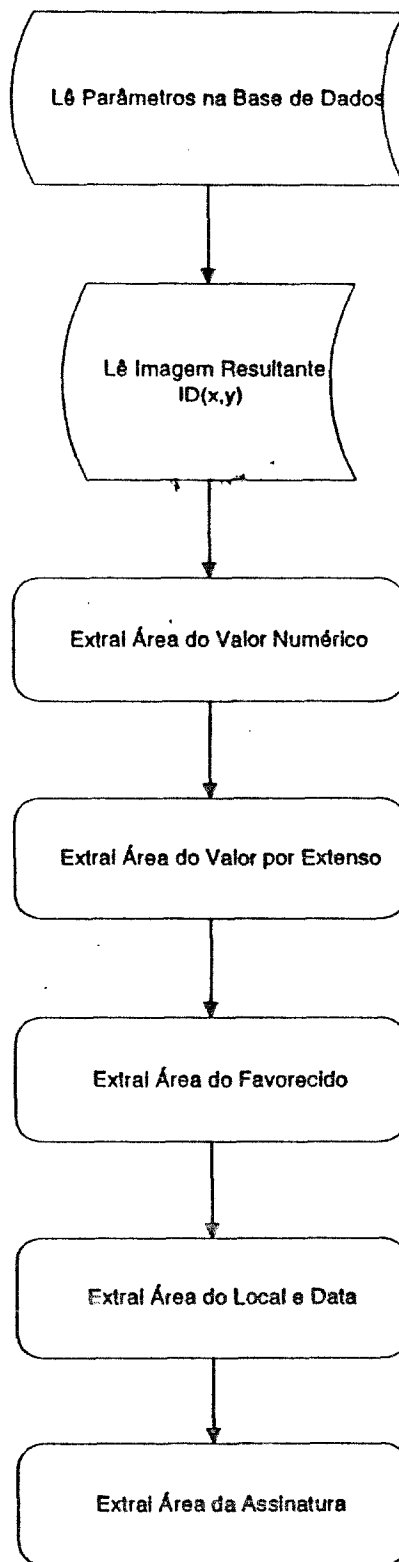
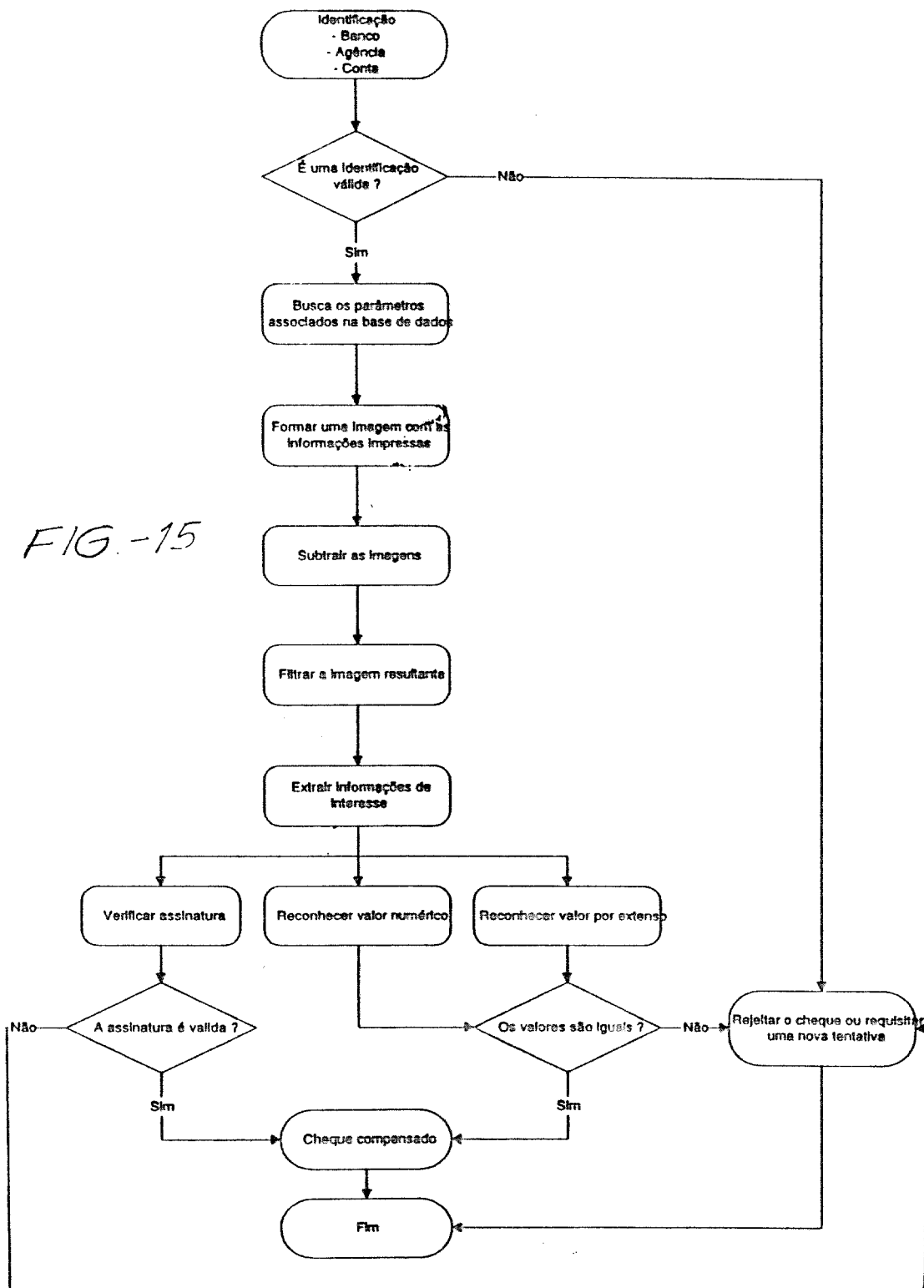


FIG.-14



456,97

Quatrocentos e cinquenta e seis reais
e noventa e sete centavos.

Quitada 01 Setembro 94
Jorge Silva

FIG.-16

456,97

Quatrocentos e cinqüenta e seis Reais
e noventa e sete centavos.

Quintila 01 Setembro 94

Coroaya Silva

FIG - 17

RESUMO

"MÉTODO E SISTEMA PARA EXTRAÇÃO AUTOMÁTICA DAS INFORMAÇÕES PREENCHIDAS EM CHEQUES BANCÁRIOS", compreendendo o emprego de um digitalizador óptico (1), uma leitora MICR (2) padrão CMC-7 e um microcomputador (3); dito método visa possibilitar a compensação automática de cheques bancários, sem a necessidade de intervenção de operadores ou digitadores.